

TORE PERSSON

**PARAMETER ESTIMATION
IN CENSORED SAMPLES
AND PROCESSES**



LUND 1984

UNIV.
LIBRARY
CAL.

Tore Persson

PARAMETER ESTIMATION
IN CENSORED SAMPLES
AND PROCESSES

1984:2

Lund 1984

TYP AV DOKUMENT ☒ Doktorsavhandling ☐ Examensarbete ☐ Konferensrapport	☐ Ansökan ☐ Reserapport ☐ Delrapport ☐ Slutrapport	☐ Tidskriftsartikel ☐ Konferenssuppsats ☐	DOKUMENTBETECKNING/CODEN LUTFD2/(TFMS-3034)/1-23/(1984)
AVDELNING/INSTITUTION Department of Mathematical Statistics Box 725 S-220 07 LUND, Sweden			
FÖRFATTARE Tore Persson			
DOKUMENTTITEL OCH UNDERTITEL Parameter estimation in censored samples and processes.			
SAMMANFATTNING Samples of independent observations or from stochastic processes can be censored in various ways. The first paper presents estimators of μ and σ for type I singly censored normal samples. Both small-sample and large-sample properties are investigated. The asymptotic efficiencies are shown to be very high. The second paper studies the exact distributions of some statistics for samples from singly truncated normal distributions. The third paper is similar to the first but considers double (two-sided) censoring. In the fourth paper the effects on the estimators in paper one and three of dependence between the observations are studied by applying the estimators to a censored stationary Gaussian process. The estimators are shown to be consistent and asymptotically normal.			
NYCKELORD Type I censoring, type II censoring, parameter estimation, normal distribution, Gaussian process.			
DOKUMENTTITEL OCH UNDERTITEL – SVENSK ÖVERSÄTTNING AV UTLÄNDSK ORIGINALTITEL Parameterskattning i censurerade stickprov och processer.			
TILLÄMPNINGSSOMRÅDE			
TILLÄMPNINGSNYCKELORD			
UTGIVNINGSDATUM år 84 mån 03	ANTAL SID 119 pages	SPRÅK ☐ svenska ☒ engelska ☐ annat	
ÖVRIGA BIBLIOGRAFISKA UPPGIFTER		ISSN	
		ISBN	
		PRIS	
I, the undersigned, being the copyright owner of the abstract, hereby grant to all reference source permission to publish and disseminate the abstract.			
Date	Signature		

1984-03-29

Tore Persson



Digitized by the Internet Archive
in 2026

CONTENTS

1.	Introduction	1
2.	Truncation and censoring principles	1
3.	Random processes	3
4.	Estimation of mean and variance	7
4.1.	Type I single censoring	9
4.2.	Type I double censoring	10
5.	Some related distributions	11
6.	Dependent observations	12
7.	Estimation of correlation coefficients	12
8.	Type II censoring	15
9.	Other distributions	19
	Appendix	21
	References	22

ACKNOWLEDGEMENTS

First I want to express my gratitude to Docent Daniel Thorburn whose detailed suggestions, encouragement, and careful criticism have led me through the work.

Further I want to thank

Professor Gunnar Blom, for valuable advice regarding the presentation of my results,

Dr Thomas B Johansson, whose problem in nuclear physics led to the first paper,

Docent Georg Lindgren, for his support and interest in my work,

Docent Holger Rootzén, for valuable cooperation on the first paper, and whose idea it was to use an observed fraction as an estimate of a true fraction.

Finally, I want to thank AB Draco for generous access to various computers, giving Ulf Lindmark for access to his graphical programme, Dagmar Jensen for excellent typing, and my family for their patience and never-failing support.

Lund, March 1984

Tore Persson



The thesis consists of this summary and the following four papers:

- [A] Persson, T. and Rootzén, H. (1977). Simple and highly efficient estimators for a type I censored normal sample. *Biometrika* 64, 123-128.
- [B] Persson, T. (1981). On the distribution of some statistics for samples of truncated normal variables. TFMS-3025. Department of Mathematical Statistics, University of Lund.
- [C] Persson, T. (1983a). Simple and highly efficient estimators for a type I doubly censored normal sample. TFMS-3030. Department of Mathematical Statistics, University of Lund.
- [D] Persson, T. (1983b). Estimating mean and variance in a clipped stationary Gaussian process. TFMS-3032. Department of Mathematical Statistics, University of Lund.

1. INTRODUCTION

Censored samples and truncated distributions provide useful statistical models in many situations. They appear for instance in medical trials, reliability theory, and quality control. During the last decades many articles and books have dealt with theoretical and applied aspects of this topic; see for instance Hald (1949), Gupta (1952), Saw (1961), Harter and Moore (1966), Nelson and Schmee (1979), Lawless (1982), and Schneider (1983). Almost all of them consider independent observations from one-dimensional distributions. Similar problems arise, however, also in connection with multivariate distributions and stochastic processes. In the thesis, truncation and censoring are considered for univariate and bivariate distributions, and for stochastic processes.

A collection of notations used is given in the appendix.

2. TRUNCATION AND CENSORING PRINCIPLES

We shall first review the usual definitions connected with censoring and truncation; cf. Hald (1949) and Gupta (1952).

Definition 2.1 A sample is *type I censored* if all values outside a fixed interval are removed.

Definition 2.2 A sample is *type II censored* if given numbers of observations in the lower and/or upper tails are removed.

Definition 2.3 A distribution (population) is *truncated* if each observation outside a given interval is rejected and replaced by a new observation, until a value inside the interval is obtained.

Definition 2.4 The truncation or type I censoring is *single* (left or right) if the interval is of the type (u_1, ∞) or $(-\infty, u_2)$, *double* if it is of the type (u_1, u_2) (u_1 and u_2 finite). The type II

censoring is *single* if values are removed from only one tail, *double* if values are removed from both tails.

The basic difference between truncation and censoring is that the former is applied to a population and the latter to a sample. The difference between truncation and type I censoring is that in the former case the effective sample size is fixed and the number of "lost" observations unknown, while in the latter case the effective sample size is random (it may even be zero) and the number of lost observations known. If a sample from an unknown population is type I censored, the number of lost observations often contains information about the population. In a type II censored sample both the effective sample size and the number of lost observations are fixed.

We shall give some examples of the various types of censoring and truncation.

Example 2.5 (Double truncation) Suppose that the output of a manufacturing process is a sequence of items whose lengths $x_1, x_2, \dots$ are independent observations from a population with distribution function F_X . All items are inspected and those who are shorter than u_1 or longer than u_2 (u_1 and u_2 fixed) are rejected. The lengths $y_1, y_2, \dots$ of the items which pass the inspection are then independent observations with the distribution function

$$F_Y(y) = \begin{cases} 0 & \text{if } y < u_1 \\ \frac{F_X(y) - F_X(u_1)}{F_X(u_2) - F_X(u_1)} & \text{if } u_1 \leq y \leq u_2 \\ 1 & \text{if } y > u_2 \end{cases}$$

Example 2.6 (Type I single censoring) A scale has a range from 0 to 120 kilogrammes. If a group of heavy people are weighed, one or more of the weights may have to be recorded only as "above 120".

Example 2.7 (Type II single censoring) In a life-testing experiment, 100 light-bulbs are switched on at the same time. It has been decided in advance to terminate the experiment when exactly ten burning bulbs remain.

Censoring and truncation can be applied to various initial distributions. Some interesting problems in connection with these operations are:

P1. Given the initial distribution, what is the distribution of the observations in an obtained sample and statistics thereof?

P2. Given a sample, how can the underlying distribution be estimated?

The mathematics involved in some of these problems is simplified if the initial distribution is (univariate) normal. During the 30's, the exact distributions of some sample statistics from censored normal samples and truncated normal distributions were studied. The best known results are those by Fisher (1931) and Geary (1935). In [B], some further results of this character are obtained. With modern computers it has also been possible to obtain more extensive numerical results.

3. RANDOM PROCESSES

The concepts of truncation and censoring can be carried over to stochastic processes in several ways. Some new problems then arise which did not appear in the univariate case.

Consider a discrete-time process $\{X_t\}_1^n$. We say that this process is type I doubly censored if all observations below a fixed level u_1 are replaced by the information "below u_1 " and every observation above

a fixed level u_2 by the information "above u_2 ".

Type I single censoring (from above or from below) is defined similarly.

Table 3.1 and Figure 3.1 show a type I doubly censored random process with $u_1 = -1$, $u_2 = 1$ and original sample size $n = 50$.

1-10	11-20	21-30	31-40	41-50
-	0.88	-0.39	-0.11	0.55
-0.91	0.76	0.24	0.28	0.59
-	0.80	0.37	-0.46	-0.75
-0.83	0.52	0.14	-0.01	-0.36
-0.57	+	0.68	0.21	0.45
-	0.86	0.91	+	-0.33
-0.48	-0.06	-0.28	+	0.13
-0.66	0.22	0.65	0.86	-0.90
0.84	0.10	-0.35	+	-0.84
0.73	-0.72	-	+	-0.68

Table 3.1 A type I doubly censored sample from a random process

with $u_1 = -1$, $u_2 = 1$, $n = 50$. (Some data as in Figure 3.1.)

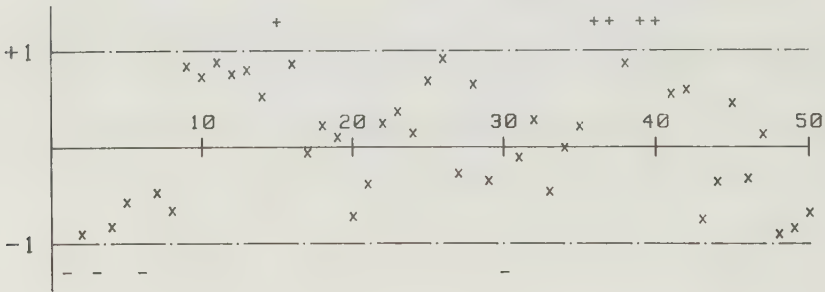


Figure 3.1 A type I doubly censored process. (Same data as in Table 3.1.)

The minus (plus) signs refer to values below u_1 (above u_2).

The definitions of type I censoring are easily extended to continuous time, and to quite general parameter sets.

Type II censoring of random processes can be defined analogously, at least if the original sample size is finite, but this kind of censoring seems to be a less natural concept, at least when the parameter

represents time. However, voluntary type II censoring may be used for, say, robustness reasons.

Truncation in the univariate case means not only to remove certain values but also to forget that they ever existed. Hence, truncation in connection with stochastic processes may be taken to mean stopping the clock as soon as a value outside the truncation points occurs, and starting it again when a value inside these points occurs, hence changing the time-scale.

Example 3.1 If a process $X_t = 1.8, 0.5, 1.4, -0.8, -1.4, 0.6, 0.8, 2.1, 0.6, -0.5$ for $t = 1, \dots, 10$ is truncated at -1 and $+1$ we get $Y_s = 0.5, -0.8, 0.6, 0.8, 0.6, -0.5$ for $s = 1, \dots, 6$.

Example 3.2 The definition of truncation can be extended to a continuous-time process $\{X_t: 0 \leq t < \infty\}$. This is illustrated in Figure 3.2.

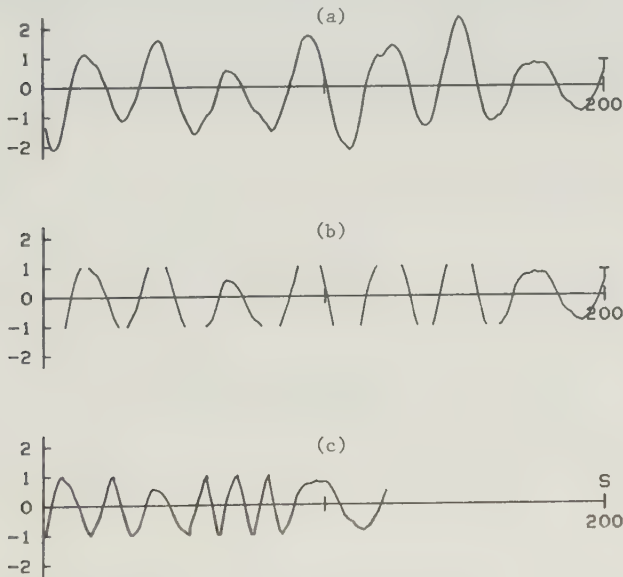


Figure 3.2 Truncation of a continuous-time process.

- (a) The initial process $\{X_t: 0 \leq t \leq 200\}$
- (b) All values outside $[-1, 1]$ have been removed
- (c) The pieces of the sample path have been pushed to the left so that no gaps remain.

Many other kinds of truncation and censoring are possible with processes. We shall give a few examples.

Example 3.3 A process $\{X_t\}$ is *clipped* at u_1 and u_2 if all values below u_1 are replaced by u_1 and all values above u_2 by u_2 . The clipped version $\{Y_t\}$ of the process is thus given by $Y_t = \max(u_1, \min(X_t, u_2))$.

Example 3.4 (Continued from Example 2.5) Assume that there is a single production line and that $x_1, x_2, \dots$ are dependent. After the inspection, $y_1, y_2, \dots$ is a truncated process. It is, however, reasonable to assume that also the total number of rejected items is known, which makes this situation intermediate between censoring and truncation.

Example 3.5 Let Y_t be the level at time t in a reservoir with capacity a and constant rate b of outflow. If the inflow is X_t during the last time unit, then Y_t is given by

$$Y_t = \min(a, \max(0, Y_{t-1} + X_t - b)) .$$

Example 3.6 A process $\{X_t\}$ is recorded by a mechanism which becomes inactive for k time-units as soon as a value above u occurs. If a value is above u during this period, it does not prolong the inactivity interval.

Example 3.7 A process $\{Y_t\}$ is obtained from $\{X_t\}$ by a linear filter. The output Y_t may be observed only if $|X_t| \leq u$. For instance, if the filter performs a simple time-shift, $Y_t = X_{t-1}$, then X_{t-1} can only be observed if $|X_t| \leq u$.

Some interesting problems in connection with censored and truncated processes are:

- P3. If the distribution of the initial process is given,
what is the distribution of the observed process?

P4. If we have observed a portion of a censored or truncated process, how can the distribution of the underlying process be estimated?

P5. How does one reconstruct the original sample path, given a censored or truncated version of it? What criteria should be used to judge the "goodness" of a reconstruction principle?

Problems P3 and P4 correspond to P1 and P2 in Section 2 while P5 is new.

The feasibility of a solution to a problem of type P3, P4, or P5 depends on the initial distribution and the truncation mechanism. We concentrate our efforts on type I censoring, in particular of stationary Gaussian processes. The situations discussed in Examples 3.3 - 3.7 will not be discussed further, but it is hoped that a thorough study of the more accessible problems will give a better insight into more complicated situations.

4. ESTIMATION OF MEAN AND VARIANCE

A natural problem given a censored sample or a sample from a truncated distribution is to estimate the underlying distribution. For instance, we may know that X is distributed as $N(\mu, \sigma^2)$, where μ and σ are unknown, or that $\{X_t\}$ is a stationary Gaussian process with unknown mean and covariance function.

Estimation in the univariate normal situation has been studied by many previous authors. For instance, Hald (1949), Halperin (1952), Gupta (1952), and Cohen (1961) have presented tables and diagrams for maximum-likelihood estimation of μ and σ , and Harter and Moore (1966) used an iterative method to solve the maximum-likelihood equations.

Estimators involving linear combinations of order statistics have been suggested for type II censored samples by Gupta (1952), Sarhan and Greenberg (1956, 1958), and the method described in Blom (1962) for producing unbiased nearly best linear estimates lends itself readily to this situation. Saw (1961) suggested several similar estimators with simple coefficients.

The estimators in paper [A] are also presented in the book by Lawless (1982), and the same idea has been adopted to double censoring by Schneider (1983).

The papers [A] and [C] deal with the problem of estimating μ and σ in a normal population given a type I singly or doubly censored sample. In [B] the distributions of some statistics involved are studied, and in [D] the influence of dependence between the observations is studied by applying the estimators derived in [A] and [C] to a censored stationary Gaussian sequence.

The basic idea is the same in [A] and [C]: an observed proportion (say, $\frac{z}{n}$) of censored values is used as an approximation to the corresponding true proportion (e.g., $\Phi(\frac{\mu-\mu}{\sigma})$). The approximation is then used to simplify the maximum likelihood equations. This leads to several explicit and relatively simple estimators of μ and σ . The estimators are found to have high asymptotic efficiencies when the censoring is severe but to be inefficient when the censoring is mild. New estimators are therefore formed by using the first ones to remove the bias from the ordinary estimators $\bar{x}$ and s . The final estimators are shown to have asymptotic efficiencies (with respect to the maximum-likelihood estimators) exceeding 98 % (99 % for some estimators) for all degrees of censoring.

Small-sample properties are studied by simulation in [A] and [C]. It is found that the suggested estimators behave almost as the corresponding maximum-likelihood estimators.

4.1. Type I single censoring

Type I single (left) censoring is considered in [A]. Let n be the initial sample size and z the number of observations which lie to the left of the censoring point u and thus are unavailable, and let $x_1, x_2, \dots, x_k$ be the remaining observations. In [A] the likelihood function is maximized under the restriction $\Phi\left(\frac{u-\mu}{\sigma}\right) = \frac{z}{n}$. In [C] the technique used is to make analogous approximations in the maximum-likelihood equations instead, but the estimators obtained in [A] could have been obtained in the same way. (On the other hand: the line of reasoning used in [A] can *not* be applied directly to double censoring.)

The maximization of the likelihood function in the singly censored case gives a pair of estimators $(\mu_{RML}^*, \sigma_{RML}^*)$ which are highly efficient if the censoring is severe, but the efficiency decreases to 0 as u tends to $-\infty$. On the other hand, $\bar{x}$ and s are good when the censoring is mild but strongly biased when u is large. We therefore use σ_{RML}^* to compensate for the bias in $\bar{x}$ and s in the following way:

$$\begin{aligned}\mu^* &= \bar{x} - \alpha^* \sigma_{RML}^* \\ \sigma^* &= \sqrt{s^2 - \{\theta^* \alpha^* - (\alpha^*)^2\} (\sigma_{RML}^*)^2}\end{aligned}\tag{4.1}$$

where

$$\begin{aligned}\sigma_{RML}^* &= \frac{\theta^* \sum_{i=1}^k (x_i - u) + \sqrt{(\theta^* \sum_{i=1}^k (x_i - u))^2 + 4k \sum_{i=1}^k (x_i - u)^2}}{2k} \\ \alpha^* &= \frac{n}{k} \phi(\theta^*) \\ \theta^* &= \Phi^{-1}\left(\frac{z}{n}\right)\end{aligned}$$

The estimators μ^* and σ^* are shown to be consistent as $n \rightarrow \infty$, and to have asymptotic relative efficiencies (with respect to the maximum-likelihood estimators) which are at least 99.6 % (for μ^*) and 98 % (for σ^*) for all parameter values.

Small-sample properties are studied by simulation in [A]. A close agreement with the maximum-likelihood estimators is found. As the degree of censoring increases, the estimators become biased but not until almost the whole sample has been removed.

4.2. Type I double censoring

In [C] double censoring at two points u_1 and u_2 is considered. As before, let n be the original sample size, and $x_1, x_2, \dots, x_k$ be the remaining observations, and let z_1 and z_2 be the numbers of unavailable observations to the left of u_1 , and to the right of u_2 , respectively. The approximations $\theta_1 = \theta_1^*$ and $-\theta_2 = \theta_2^*$, where $\theta_1^* = \frac{u_1 - \mu}{\sigma}$, $\theta_1^* = \Phi^{-1}(\frac{z_1}{n})$, and $\theta_2^* = \Phi^{-1}(1 - \frac{z_2}{n})$, can be introduced in several ways into the maximum-likelihood equations. The simplest way is not to use the maximum-likelihood equations at all but only the initial approximations, thus obtaining the estimators

$$\mu_q^* = \frac{u_1 \theta_2^* - u_2 \theta_1^*}{\theta_2^* - \theta_1^*}, \quad \sigma_q^* = \frac{u_2 - u_1}{\theta_2^* - \theta_1^*} \quad (4.2)$$

These "quantile estimators" were studied already by Saw (1961). Their efficiencies are surprisingly high if the censoring is severe.

Several ways of using θ_1^* and θ_2^* in the maximum-likelihood equations are investigated in [C]. The final choice (for its efficiency rather than its simplicity) yields, after combination with $\bar{x}$ and s , the estimators μ_2^{**} and σ_2^{**} below.

$$\begin{aligned} \mu_2^{**} &= \bar{x} - \frac{n}{k} (\varphi(\theta_1^*) \sigma_{22}^* - \varphi(\theta_1^*) \sigma_{21}^*) \\ \sigma_2^{**} &= \sqrt{s^2 - \left\{ \frac{n}{k} (\theta_1^* \varphi(\theta_1^*) - \theta_2^* \varphi(\theta_2^*)) - \left(\frac{n}{k} (\varphi(\theta_1^*) - \varphi(\theta_2^*)) \right)^2 \right\} (\sigma_2^*)^2} \end{aligned} \quad (4.3)$$

where

$$\sigma_{21}^* = \frac{1}{2} \{A_1 + \sqrt{A_1^2 + \frac{4}{k} \sum_1^k (x_1 - u_2)^2}\}$$

$$\sigma_{22}^* = \frac{1}{2} \{A_2 + \sqrt{A_2^2 + \frac{4}{k} \sum_1^k (x_1 - u_1)^2}\}$$

$$A_1 = \frac{n}{k} (u_2 - u_1) \varphi(\theta_1^*) + \theta_2^* (\bar{x} - u_2)$$

$$A_2 = \frac{n}{k} (u_2 - u_1) \varphi(\theta_2^*) + \theta_1^* (\bar{x} - u_1)$$

$$\sigma_2^* = \frac{\varphi(\theta_2^*) \sigma_{21}^* + \varphi(\theta_1^*) \sigma_{22}^*}{\varphi(\theta_1^*) + \varphi(\theta_2^*)}$$

The asymptotic relative efficiencies of μ_2^{**} and σ_2^{**} (with respect to the maximum-likelihood estimators) are at least 99.4 % (for μ_2^{**}) and 99.2 % (for σ_2^{**}), respectively, for all values of θ_1 and θ_2 .

In [C] small-sample properties of the proposed estimators are studied by simulation. As in 4.1 a close agreement with the corresponding maximum-likelihood estimators is found.

5. SOME RELATED DISTRIBUTIONS

One of the reasons for studying the small-sample properties in [A] and [C] by simulation was that the exact distribution of $\left(\sum_1^k X_1, \sum_1^k X_1^2 \right)$ is not known, if $X_1, X_2, \dots, X_k$ are independent random variables with a truncated normal distribution. All estimators studied in [A] and [C] are, essentially, functions of this statistic.

In paper [B] the distributions of the statistic $\left(\sum_1^k X_1, \left(\sum_1^k X_1^2 \right)^{\frac{1}{2}} \right)$ and some functions of it when $X_1, X_2, \dots, X_k$ are singly truncated, are derived analytically for $k = 2, 3, 4$, numerically for $k \leq 20$, and asymptotically as $k \rightarrow \infty$.

A number of related results are also obtained in [B]. Perhaps the most noteworthy of these is a generalization of the well-known fact that the sum of the squares of k independent $N(\mu, \sigma^2)$ variables is a

stochastic mixture of χ^2 distributions: this property carries over in a natural way to singly truncated normal variables.

6. DEPENDENT OBSERVATIONS

In paper [D] the estimators μ^* and σ^* given in [A] and [C] are applied to a type I singly censored stationary Gaussian sequence $\{X_t\}_1^n$. It is shown that the estimators are consistent as $n \rightarrow \infty$, provided that $\sum_{\tau=1}^{\infty} r(\tau) < \infty$ and $\sum_{\tau=1}^{\infty} r^2(\tau) < \infty$, where r is the covariance function of $\{X_t\}$. It is also shown by use of a theorem of Cuzick (1976) that the joint asymptotic distribution of (μ^*, σ^*) is bivariate normal, if $\sum_{\tau=1}^{\infty} |r(\tau)| < \infty$.

Some small-sample properties are investigated in [D] by simulation with sample sizes up to $n = 100$ for a number of AR(1) processes. The main impression is that the bias is larger than in the independent case if $\rho(X_t, X_{t+1})$ is close to $+1$ or -1 . This is seen most clearly when the censoring is severe.

7. ESTIMATION OF CORRELATION COEFFICIENTS

A problem which is closely related to the problems discussed in [D] is that of estimating the covariance function of a censored stationary Gaussian process.

The usual estimator is given by $r^*(\tau) = \frac{1}{k} \sum (X_t - \bar{X})(X_{t+\tau} - \bar{X})$ where the summation is extended over all complete pairs $(X_t, X_{t+\tau})$. The denominator may be chosen to be $k = n - \tau$ or $k =$ the number of complete pairs. Neither of these gives a consistent estimator of $r(\tau)$ as $n \rightarrow \infty$, nor is $r^*(\tau)/r^*(0)$ a consistent estimator of the correlation coefficient $\rho(\tau)$.

A similar, and simpler, problem is that of estimating the correlation coefficient ρ in a censored sample from a bivariate normal distribution.

A rather detailed study of this problem is given here since it has not been studied elsewhere.

We say that a sample $\{(x_i, y_i)\}_1^n$ from a bivariate distribution is type I censored from below at (u_1, u_2) if only those x_i which are above u_1 and those y_i which are above u_2 can be observed. There are, say, n_{11} pairs where neither x_i nor y_i is available, n_{12} pairs where y_i is available but not x_i , n_{21} pairs where x_i is available but not y_i , and $n_{22} = n - n_{11} - n_{12} - n_{21}$ pairs, where $x_i > u_1$ and $y_i > u_2$ and both thus available.

If the initial distribution is $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then ρ can be estimated as follows: The expected values of $n_{11} + n_{12}$, $n_{11} + n_{21}$, and n_{11} are $\Phi(\theta_1)$, $\Phi(\theta_2)$, and $\Phi_2(\theta_1, \theta_2; \rho)$, respectively, where $\theta_i = \frac{u_i - \mu_i}{\sigma_i}$. Estimates of θ_1 , θ_2 , and ρ may be obtained by solving the following set of equations.

$$\left. \begin{aligned} \Phi(\theta_1^*) &= \frac{n_{11} + n_{12}}{n} \\ \Phi(\theta_2^*) &= \frac{n_{11} + n_{21}}{n} \\ \Phi_2(\theta_1^*, \theta_2^*; \rho^*) &= \frac{n_{11}}{n} \end{aligned} \right\} \quad (7.1)$$

It can be shown that (7.1) always has at least one solution if $\theta_i^* = \pm\infty$ are allowed, and that the solution is unique if both θ_1^* and θ_2^* are finite (i.e., if neither $n_{11} + n_{12}$ nor $n_{11} + n_{21}$ is zero or n).

For:

$$\Phi_2(\theta_1^*, \theta_2^*; -1) = \begin{cases} 0 & \text{if } \theta_1^* + \theta_2^* \leq 0 \\ \Phi(\theta_1^*) + \Phi(\theta_2^*) - 1 & \text{if } \theta_1^* + \theta_2^* > 0 \end{cases}$$

$$\Phi_2(\theta_1^*, \theta_2^*; +1) = \Phi(\min(\theta_1^*, \theta_2^*)) = \min(\Phi(\theta_1^*), \Phi(\theta_2^*))$$

If θ_1^* and θ_2^* satisfy (7.1), then

$$\Phi(\theta_1^*) + \Phi(\theta_2^*) - 1 = \frac{2n_{11} + n_{12} + n_{21}}{n} - 1 = \frac{n_{11} - n_{22}}{n}$$

$$\min(\phi(\theta_1^*), \phi(\theta_2^*)) = \frac{n_{11} + \min(n_{12}, n_{21})}{n}$$

Thus,

$$\phi_2(\theta_1^*, \theta_2^*; -1) \leq \frac{n_{11}}{n} \leq \phi_2(\theta_1^*, \theta_2^*; +1)$$

which proves the existence part of the statement, since ϕ_2 is continuous.

The uniqueness follows from the fact that $\frac{\partial}{\partial \rho} \phi_2(\theta_1^*, \theta_2^*; \rho) = \phi_2(\theta_1^*, \theta_2^*; \rho) > 0$ for finite θ_1^* and θ_2^* .

A special case of particular interest is when the distribution and the censoring are symmetric ($\mu_1 = \mu_2$, $\sigma_1 = \sigma_2$, $u_1 = u_2$). Instead of (7.1) it is then natural to solve

$$\left. \begin{aligned} \phi(\theta^*) &= \frac{2n_{11} + n_{12} + n_{21}}{2n} \\ \phi(\theta^*, \theta^*; \rho^*) &= \frac{n_{11}}{n} \end{aligned} \right\} \quad (7.2)$$

This system of equations also has at least one solution, and exactly one if neither n_{11} nor n_{22} is equal to n .

The solutions to (7.1) and (7.2) give consistent estimators of θ_1 , θ_2 , θ , and ρ when $n \rightarrow \infty$. Their efficiencies are not very high. The asymptotic variance of ρ^* can be found by expanding the left-hand members of (7.1) and (7.2) in Taylor series around the true parameter values. For instance, in the symmetric case,

$$\begin{aligned} n \cdot \text{Var}(\rho^*) &\rightarrow \frac{\phi^2(\theta)(1-\phi(\theta))^2}{\phi^4(\theta)} \quad \text{if } \rho = 0 \\ &\rightarrow \frac{\phi_2^2(0,0;\rho)(1-2\phi_2(0,0;\rho))}{2\phi_2^2(0,0;\rho)} = \left\{ \frac{\pi^2}{4} - (\arcsin \rho)^2 \right\} (1-\rho^2) \\ &\quad \text{if } \theta = 0 \\ &\rightarrow \frac{\pi^2}{4} \quad \text{if } \rho = \theta = 0 \end{aligned}$$

Since $\text{Var}(\hat{\rho}) = \frac{1}{n}$ asymptotically for the maximum-likelihood estimator $\hat{\rho}$ when $\rho = \theta = 0$, this means that the asymptotic efficiency of ρ^* is $\frac{4}{\pi^2} \approx 0.41$ in this case.

As in 4.1 and 4.2 a substantial improvement could probably be achieved by using ρ^* to remove the bias in the correlation coefficient r based on the n_{22} complete pairs. (The bias is $(I_{11}I_{00} - I_{10}^2)/(I_{20}I_{00} - I_{10}^2) - \rho$ where the I_{ab} 's can be found in the appendix of [D] with $u = \theta$.) The ordinary correlation coefficient in a non-censored normal sample is, however, not an efficient estimator of ρ unless $\rho = 0$. Therefore a combination of r and ρ^* can not be expected to give as good results as in 4.1 and 4.2.

If we have a type I singly (from below) censored stationary Gaussian sequence $\{X_t\}_1^n$, the correlations $\rho(\tau)$ of the underlying process can be estimated by applying the solution to (7.2) to the pairs $\{(X_t, X_{t+\tau})\}_{t=1}^{n-\tau}$. Then $\rho^*(\tau)$ is a consistent estimator of $\rho(\tau)$ as $n \rightarrow \infty$, if $\rho(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$. Since the pairs $(X_t, X_{t+\tau})$ are dependent, little is known about the variance of $\rho^*(\tau)$.

If the censored process is an ARMA process, there is a direct relationship between the covariance function and the parameters $\{a_i\}_1^p$ and $\{b_i\}_1^q$. It is thus possible to obtain consistent estimators of the parameters in a censored ARMA process.

If a bivariate sample or stationary sequence is doubly censored, several estimators of the ρ^* type above can be constructed by collapsing categories, i.e., by considering only one censoring point for each coordinate at a time.

8. TYPE II CENSORING

All the estimators discussed in [A], [C], and Section 7 can be applied to type II censored samples by simple changes in the meaning of u_i and θ_i^* . The asymptotic properties do not change. If we have a singly censored sample of original size n where the leftmost np observations have been removed, the estimators (4.1) can be applied if we replace θ^*

by $\Phi^{-1}(p) = -\lambda_p$ and u by x_L , the smallest remaining observation.

We thus take

$$\begin{aligned}\mu^* &= \bar{x} - \frac{\varphi(\lambda_p)}{1-p} \sigma_{RML}^* \\ \sigma^* &= \sqrt{s^2 + \left\{ \frac{\lambda_p \varphi(\lambda_p)}{1-p} + \left(-\frac{\varphi(\lambda_p)}{1-p} \right)^2 \right\} (\sigma_{RML}^*)^2}\end{aligned}\quad (8.1)$$

where

$$\sigma_{RML}^* = \frac{1}{2} \left\{ -\lambda_p (\bar{x} - x_L) + \sqrt{\lambda_p^2 (\bar{x} - x_L)^2 + \frac{4}{k} \sum_1^k (x_i - x_L)^2} \right\}$$

$$k = n - np$$

$x_1, \dots, x_k$ are the remaining observations

If a sample is doubly censored by removal of the np_1 leftmost and np_2 rightmost observations, we can use (4.3) by replacing θ_1^* by $-\lambda_{p_1}$, θ_2^* by λ_{p_2} , u_1 by x_L , and u_2 by x_R (the smallest and largest remaining observation). In particular, if the censoring is symmetric, $p_1 = p_2 = p$, say, we have

$$\begin{aligned}\mu_{22}^{**} &= \bar{x} - \frac{\varphi(\lambda_p)}{1-2p} (\sigma_{22}^* - \sigma_{21}^*) \\ \sigma_{22}^{**} &= \sqrt{s^2 + \frac{2\lambda_p \varphi(\lambda_p)}{1-2p} (\sigma_2^*)^2}\end{aligned}\quad (8.2)$$

where

$$\begin{aligned}\sigma_2^* &= \frac{\sigma_{21}^* + \sigma_{22}^*}{2} \\ \sigma_{21}^* &= \frac{1}{2} \left\{ A_1 + \sqrt{A_1^2 + \frac{4}{k} \sum_1^k (x_i - x_R)^2} \right\} \\ \sigma_{22}^* &= \frac{1}{2} \left\{ A_2 + \sqrt{A_2^2 + \frac{4}{k} \sum_1^k (x_i - x_L)^2} \right\} \\ A_1 &= \frac{\varphi(\lambda_p)}{1-2p} (x_R - x_L) + \lambda_p (\bar{x} - x_R) \\ A_2 &= \frac{\varphi(\lambda_p)}{1-2p} (x_R - x_L) - \lambda_p (\bar{x} - x_L)\end{aligned}$$

$$k = n - 2np$$

$x_1, \dots, x_k$ are the remaining observations

The estimators (8.1) and (8.2) offer robustness against outliers (onesided in the former case) if a non-censored sample is censored voluntarily.

Below we make comparisons between $(\mu_2^{**}, \sigma_2^{**})$ and some well-known estimators of μ and σ in type II censored samples: the trimmed and Winsorized means and (scaled) standard deviations. We also include the type II counterpart of (4.2),

$$\mu_q^* = \frac{x_L + x_R}{2}, \quad \sigma_q^* = \frac{x_R - x_L}{2\lambda_p} \quad (8.3)$$

and the usual (full-sample) mean and standard deviation. All estimators (except $\bar{x}$ and s) are computed with the same fractions of censored values at each end.

Figures 8.1 and 8.2 show the influence functions of the estimators for an underlying normal distribution. The influence function of a statistic T describes how much T is affected when the population is contaminated by a small point mass. Formally, the influence function $IC(x, F, T)$ of a statistic T at a distribution F is defined as the derivative $\frac{dT}{dt}$ when the distribution of the population is changed from F to $(1-t)F + t\delta_x$, where δ_x is a point mass at x . (For a detailed description of influence functions, trimming and Winsorization, see Hampel (1971) or Huber (1981).)

As can be seen from Figures 8.1 and 8.2, the influence functions of μ_2^{**} and σ_2^{**} are very close to those of the corresponding Winsorized estimators.

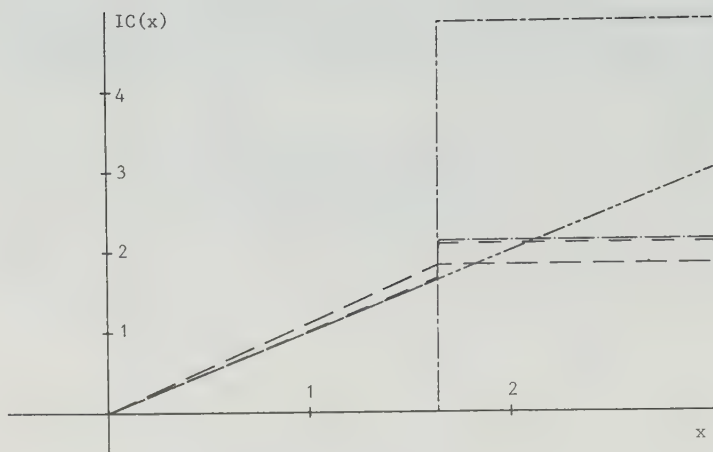


Figure 8.1 Influence functions for μ_2^{**} (---), μ_q^* (— · —), trimmed mean (— — —), Winsorized mean (— — — —), and usual mean (· · · · ·). The fraction of censored values is 0.05 at each end for the first four. The distribution is $N(0,1)$.

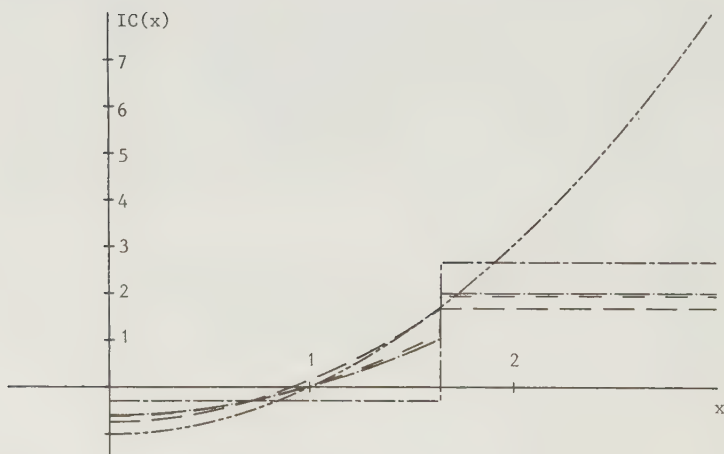


Figure 8.2 Influence functions for σ_2^{**} (---), σ_q^* (— · —), trimmed standard deviation (— — —), Winsorized standard deviation (— — — —), and usual standard deviation (· · · · ·). The same distribution and censoring as in Figure 8.1.

9. OTHER DISTRIBUTIONS

The idea of using observed proportions as estimators of true proportions and thus obtain approximate, explicit solutions to the maximum-likelihood equation(s) is not limited to normal distributions. It is, however, not always obvious *how* the estimated proportions should be used. Below some estimators are suggested. No attempt has been done to study their properties in any detail.

If the underlying distribution function is of the form $F(\frac{x-\mu}{\sigma})$ where F is known but the location and scale parameters μ and σ unknown some of the previous methods may be applied to a type I or type II censored sample.

1. The quantile estimators μ_q^* and σ_q^* in (4.2) can be used if θ_1^* and θ_2^* are defined by $F(\theta_1^*) = \frac{z_1}{n}$ and $1 - F(\theta_2^*) = \frac{z_2}{n}$ in a type I censored sample. If type II censoring is used, the estimators become

$$\mu_q^* = \frac{x_L F^{-1}(1-p_2) - x_R F^{-1}(p_1)}{F^{-1}(1-p_2) - F^{-1}(p_1)} ; \sigma_q^* = \frac{x_R - x_L}{F^{-1}(1-p_2) - F^{-1}(p_1)}$$

In both cases, μ_q^* and σ_q^* are consistent estimators of μ and σ .

2. If F is twice differentiable, the general counterpart of σ_{11}^* and σ_{12}^* in [C] can be derived (here presented in the type I case):

$$\sigma_{11}^* = f(\theta_2^*) \frac{u_2 - u_1}{k/n} + \frac{1}{k} \sum_1^k \left(\frac{\theta_2^*(x_i - u_1) - \theta_1^*(x_i - u_2)}{u_2 - u_1} \right) \cdot \frac{x_i - u_2}{u_2 - u_1}$$

$$\sigma_{12}^* = f(\theta_1^*) \frac{u_2 - u_1}{k/n} + \frac{1}{k} \sum_1^k \left(\frac{\theta_2^*(x_i - u_1) - \theta_1^*(x_i - u_2)}{u_2 - u_1} \right) \cdot \frac{x_i - u_1}{u_2 - u_1}$$

where $f(x) = e^{-Q(x)}$ is the probability density function $F'(x)$. Several combinations of these are possible, and

$\frac{\sigma_{11}^* + \sigma_{12}^*}{2}$ is not necessarily the best. The location parameter μ may be estimated by $u_1 - \theta_1^* \sigma^*$, $u_2 - \theta_2^* \sigma^*$, or combinations thereof. All the estimators mentioned are consistent. As before, a type II censored sample can be treated similarly.

3. There is in general no counterpart to μ_2^{**} and σ_2^{**} .

We finally give some examples where the estimators above may be of interest.

Example 9.1 (Extreme value distribution) Assume that we have a doubly censored sample from a distribution function of the type $x \rightarrow \exp(-\exp(-\frac{x-a}{b}))$, where a and b are unknown. The estimators above can be applied directly.

Example 9.2 (Weibull) If the distribution function of X is of the type $F_X(x) = 1 - \exp(-(\frac{x}{d})^c)$, then $-\ln X$ has an extreme value distribution with $b = \frac{1}{c}$, $a = -\ln d$ and the estimators above may thus be used together with a transformation.

No details are known about any asymptotic or small-sample properties with these other distributions, except consistency.

APPENDIX

The following notations are used in the summary.

The normal distribution with mean μ and variance σ^2 is denoted by $N(\mu, \sigma^2)$.

The density and distribution functions of $N(0, 1)$, the standardized normal distribution, are given by

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \text{and} \quad \Phi(x) = \int_{-\infty}^x \varphi(t) dt$$

The upper p^{th} quantile of $N(0, 1)$ is denoted by

$$\lambda_p = \Phi^{-1}(1-p), \text{ i.e., } \Phi(\lambda_p) = 1 - p$$

The bivariate normal distribution is denoted by $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

The density and distribution functions of $N(0, 0, 1, 1, \rho)$ are denoted by

$$\varphi_2(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right)$$

and

$$\Phi_2(x, y; \rho) = \int_{-\infty}^x \int_{-\infty}^y \varphi_2(s, t; \rho) ds dt$$

REFERENCES

- Blom, G. (1962). Nearly best linear estimates of location and scale parameters. In: Sarhan and Greenberg. Contributions to Order Statistics. John Wiley and Sons, New York, 34-46.
- Cohen, A.C. (1961). Tables for maximum-likelihood estimates: singly truncated and singly censored samples. *Technometrics* 3, 535-541.
- Cuzick, J. (1976). A central limit theorem for the number of zeroes of a stationary Gaussian process. *Ann. Probability* 4, 547-556.
- Fisher, R.A. (1931). Tables of the Hh functions. In: Mathematical tables, Volume I. British Association for the Advancement of Science, London, xxiv-xxxv.
- Geary, R.C. (1935). The ratio of the mean deviation to the standard deviation as a test of normality. *Biometrika* 27, 310-332.
- Gupta, A.K. (1952). Estimation of the mean and standard deviation of a normal population from a censored sample. *Biometrika* 39, 260-273.
- Hald, A. (1949). Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point. *Skand. Aktuar.* 32, 119-134.
- Halperin, M. (1952). Estimation in the truncated normal distribution. *J. Am. Statist. Assoc.* 47, 457-465.
- Hampel, F. (1971). A general qualitative definition of robustness. *Ann. Math. Statist.* 42, 1887-1896.
- Harter, H.L. and Moore, A.H. (1966). Iterative maximum-likelihood estimation of the parameters of normal populations from singly and doubly censored samples. *Biometrika* 53, 205-213.
- Huber, P.J. (1981). *Robust Statistics*. John Wiley and Sons, New York, 13, 57-59, 112-113.
- Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, New York, 225-226.

- Nelson, W. and Schmee, J. (1979). Inference for (log) normal life distributions from small singly censored samples and BLUE's. *Technometrics* 21, 43-54.
- Sarhan, A.E. and Greenberg, B.G. (1956). Estimation of location and scale parameters by order statistics from singly and doubly censored samples, Part I. *Ann. Math. Statist.* 27, 427-451.
- Sarhan, A.E. and Greenberg, B.G. (1958). Estimation of location and scale parameters by order statistics from singly and doubly censored samples, Part II. *Ann. Math. Statist.* 29, 79-105.
- Saw, J.G. (1961). Estimation of the normal population parameters given a type I censored sample. *Biometrika* 48, 367-377.
- Schneider, H. (1983). Simple and highly efficient estimators for censored normal samples. Technical Report. University of North Carolina, Chapel Hill.

